AD_____

Award Number: W81XWH-11-1-0261

TITLE: Use of eQTL Analysis for the Discovery of Target Genes Identified by GWAS

PRINCIPAL INVESTIGATOR: Dr. Stephen Thibodeau

CONTRACTING ORGANIZATION: Mayo Clinic
Rochester, MN 55905

REPORT DATE: April 2012

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
01-04-2012	Annual	1 APR 2011 - 31 MAR 2012

4. TITLE AND SUBTITLE
Use of eQTL Analysis for the Discovery of Target Genes Identified by GWAS

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-11-1-0261

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Dr. Stephen Thibodeau

E-Mail: sthibodeau@mayo.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Mayo Clinic
Rochester, MN 55905

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The goals of this grant proposal are to: 1) construct a prostate tissue-specific expression quantitative trait loci (eQTL) dataset; and 2) utilize this dataset to identify candidate genes for existing prostate cancer (PC) risk-single nucleotide polymorphisms (SNPs) that can then be followed up in future studies. To accomplish this goal, we will perform a genome-wide SNP analysis (Illumina Human Omni 2.5M SNP array) and a genome-wide mRNA expression analysis (Illumina humanht-12 BeadChip) on a common set of 500 samples of normal prostate tissue sampled from men with PC. To date, we have pre-screened normal prostate tissue with the use of H&E stained sections from 4000 men having a radical prostatectomy in order to identify those cases meeting our strict selection criteria for further processing (tissue localized to the posterior region of the prostate, no tumor, no high grade PIN, no BPH, <1% lymphocytes, and the final percent of epithelial cells present >40%). Following this initial evaluation, 565 tissue samples were selected and further processed into 10 micron thick sections (cryostat) and reevaluated (H&E stained sections) prior to final processing for RNA and DNA extraction. Of the 565, 524 samples have been selected for RNA and DNA extraction, which is now in progress and should be completed shortly.

15. SUBJECT TERMS
eQTL dataset

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	6	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

Page

A. INTRODUCTION:

We hypothesize that many of the PC disease-associated SNPs already identified to date will be located in regulatory domains involved in gene transcription. Furthermore, we hypothesize that candidate genes affected by these regulatory elements can be identified by taking advantage of eQTL datasets. Therefore, the objectives of this grant proposal are to: 1) construct a prostate tissue-specific eQTL dataset that can be used to identify candidate genes for any current (or future), predictive (or prognostic) SNP identified for PC; and 2) utilize this dataset to identify candidate genes for existing PC risk SNPs that can then be followed up in future studies. To accomplish this goal, we proposed to perform a genome-wide SNP analysis (Illumina Human Omni 2.5M SNP array) and a genome-wide mRNA expression analysis (Illumina humanht-12 BeadChip) on a common set of 500 samples of normal prostate tissue sampled from men with PC. The long-term objective of this strategy is to characterize the functional role of the disease-causing SNPs, to identify the biologic pathways affected by these inherited factors, and ultimately to identify targets for disease prediction, risk stratification and identification of treatment targets.

B. BODY:

Statement of work originally proposed for year 1:

Task 1. Processing of normal prostate tissue for RNA purification (months 1-9)

- **1a.** Cryo-section fresh-frozen tissue from ~500-600 cases (months 1-9)
- **1b.** Create hematoxylin-eosin stained slides from each case for review (months 1-9)
- **1c.** Review of sections by a Pathologist. (months 1-9).
- **1d.** Select 500 cases of high-quality samples for RNA extraction (Task 2) (months 10).

Task 2. RNA Extraction from 500 cases for study (months 11-12)

- **2a.** Use sections from 500 samples selected from Task 1 to purify total RNA (months 11-12)

Task 3. Genome-wide genotyping of blood DNA from 500 cases for study (months 12-14)

- **3a.** Place blood DNA (already extracted) in 96 well plates for genotyping (months 12)
- **3b.** Genotype samples (months 12-14)
- **3c.** Quality-control checks and data processing – Statistical analyses (months 14)

Work performed: Task 1 (Processing of normal prostate tissue for RNA purification)

All of the work proposed for Task 1 has now been completed. In order to achieve our goal of 500 samples of normal prostate tissue, we initially reviewed H&E stained sections from all archived cases available for study; ~4,000. These ~4000 cases were obtained from patients whom had undergone a radical prostatectomy at Mayo Clinic and are available to investigators through the Prostate Cancer SPORE. Typically, one to three pieces of frozen tissue (snap frozen at the time of surgery) is available for each case. At the time each case is initially processed, a representative H&E stained slide is made from each piece of tissue and archived for future investigator review to aid in the process of tissue selection. Although the archived slide allows for an initial evaluation, blocks are used over time and the histology can change. Thus, cutting an additional representative H&E is often necessary to re-evaluate the current state of these blocks.

For this study, the same pathologist was used throughout the evaluation process to ensure consistency. In our initial pre-screen of the ~4000 normal tissue cases, we first removed all cases where the patient's tumor had a Gleason score greater than 7, cases where tumor was found on the H&E slide and cases where normal prostate tissue was not available. Following this initial review, 916 pieces of tissue were available for further processing. The archived tissue was then pulled from long-term storage and a fresh representative H&E stained slide was prepared for re-evaluation by a Pathologist. In order to meet the needs of this study, the following criteria were developed for further tissue selection and processing:

1. No tumor present on the new H&E.
2. The section viewed had to be from the posterior region of the prostate – all central and anterior zone tissues were eliminated. The region of interest was determined based on histologic landmarks and Mayo practice processes (posterior region are inked for orientation).
3. No High Grade Prostatic Intraepithelial Neoplasia (HGPIN).
4. No greater than 1% of the cells on the slide could be lymphocytes.
5. The final percent of epithelial glands present on the slide had to be at least 40%.

Of the 916 cases re-examined, 93 cases met the criteria above, but also contained Benign Prostatic Hyperplasia (BPH), seminal vesicle, urethra, or adjacent central zone. These pieces of tissue were further processed to eliminate the contaminating portion and an additional H&E stained section was prepared to ensure that the block was processed correctly and the unwanted regions were adequately removed.

Following the final review of tissue, 565 cases met the selection criteria noted above. Due to the small number of cases meeting our strict histologic criteria (565 of ~4000 cases reviewed), most of the selected cases did not have blood available for the extraction of DNA (for genotyping). As a result, we chose to take additional sections of the normal prostate tissue, which allowed for the extraction of both RNA (expression) and DNA (genotyping). From past experience, we expected that a degree of histologic change would be present throughout the sectioning process and this would results in an additional ~10% of the cases failing to meet our selection criteria. Thus, we decided to section and evaluate all 565 cases, re-evaluate H&E stained sections once more and then choose the best cases for the final processing.

For the extraction of DNA and RNA, tissue was first sectioned on a cryostat, preparing 10-micron thick sections. Prior to sectioning, however, all of the samples were randomized into cutting groups based on percent epithelium, presence or absence of lymphocytes, the time of original tissue collection, and if the tissue came from prostate cancer patients or from patients having a cysto-prostatectomy due to bladder cancer. The randomization of samples was performed in order to control for any cutting bias that might be introduced as the tissue was processed each day. The 565 cases were sectioned over a period of 26 working days in the following manner: the initial section was taken for an H&E stained slide (to serve in a one-to-one comparison with the initially reviewed H&E section to confirm that no tissue mix-up had occurred), then multiple sections placed in tube 1 for RNA, a 2nd H&E section, multiple sections placed in tube 2 for RNA, 3rd H&E section, multiple sections placed in tube 3 for DNA, 4th H&E sections, multiple sections placed in tube 4 for DNA, and the final H&E section. For the RNA destined tubes, tissue was immediately placed in QIAzol buffer and then snap frozen to ensure high-quality RNA. For the DNA destined tubes, sections were placed in tubes and initially stored at -80 C. These tubes were then collected the following day, and cell lysis buffer and proteinase K were added to both DNA tubes and digested overnight at 55 C on a shaking incubator essentially as outlined by the manufacturer. Visual confirmation was done the following day to ensure all of the tissue was digested, and then the tubes were considered stable and stored at 4 C pending completion of the DNA extraction.

All five H&Es sections outlined above were evaluated once again by a Pathologist to ensure that no histologic changes had occurred as the tissue was sectioned. Additionally, the 1st H&E was used to compare to the original H&E confirming that no specimen mix-ups had occurred. Upon histologic review of all five H&E slides, roughly 10% of the cases were eliminated due to histologic changes (i.e. the appearance of small cancer foci, change in % epithelium, appearance of HGPIN, an increase in lymphocytic presence) as predicted. Following this final review, 505 cases remained that met the initial criteria. Again, because we anticipate that there would be a small number of cases that have poor-quality RNA or poor DNA yield, an additional 19 cases were selected that had 2% infiltrative lymphocytes present for the final process of DNA and RNA extracted. These 524 cases were then split into two batches for RNA extraction and re-randomized again as previously described, but now the randomization scheme also including the day the tissue was processed. This randomization was performed to avoid any batch effects during RNA extraction.

Work performed: Task 2 (RNA Extraction from 500 cases for study)

The extraction of both tissue RNA and DNA is currently underway, with a completion date anticipated by the end of April. The completion of the DNA extraction will coincide with the RNA extraction completion time. Once completed, we will select the optimum 500 samples for the mRNA expression and genotyping studies based on RNA and DNA quality and those meeting the most strict selection criteria (i.e. higher % epithelium, no or fewest lymphocytes present).

C. KEY RESEARCH ACCOMPLISHMENTS:

- Tissue processing for the construction of the eQTL data set is now complete.
- The extraction of tissue RNA and DNA is currently in progress and will be completed shortly.

D. REPORTABLE OUTCOMES:

- In the process of construction an eQTL dataset.

E. CONCLUSION:

The first major goal of this proposal is to construct a prostate tissue-specific expression quantitative trait loci (eQTL) dataset. Tissue processing for the construction of this eQTL data set has now been completed. The extraction of tissue RNA and DNA is currently in progress and will be completed shortly. Once completed, we will then be in a position to complete the mRNA expression and DNA genotyping studies.

F. REFERENCES: None

G. APPENDICES: None